

対戦形知能ロボット制御用先読み探索プロセッサの構成法

苔米地 宣 裕*・藤 枝 孝 徳**

Design of the Search Processor for Control of Game Playing Intelligent-Robots

Nobuhiro TOMABECHI* and Takanori FUJIEDA**

Abstract

This paper presents a design of the search processor for control use of game playing intelligent-robots. The game model is assumed that two robots autonomously travel around the objects distributed in an area and competitively collect the objects. The features of the design are as follows; ① A search algorithm suitable for the game model named the un-alternatively played α - β method is employed. ② The search functions are realized by the hardware technology as well as possible. ③ A multi-processor system is employed in which multiple processors built in a VLSI chip share a search job. It is demonstrated that the searching speed of the proposed processor is 42 times that of the conventional computer for a single processor.

Key words : search, processor, game, intelligent, robot

1. ま え が き

本研究では、自分と利益が反する相手があり、相手と競争しながら何かを実行するような種類のロボット（対戦ロボットという）を対象とする。対戦ロボット制御に特徴的な課題は、動力学的な問題よりも、むしろ、対戦に勝利するための知的な能力をいかに付与するかにある。これは、優れて人工知能的課題^{(1),(2)}であると言える。対戦に勝利するための知的な能力には、様々なものが考えられるが、本研究では、先読み能力について取り扱う。一般に、先読みは、読みの深さが深くなると処理時間が飛躍的に大きくなる。しかも、対戦ロボットには、リアルタイム応答が要求される。

本論文では、いろいろな種類の対戦ロボットに共通に使用でき、かつ、高速性を有する先読

み探索プロセッサの実現を目標とする。先読みは、あらゆる対戦に共通するが、しかし、プロセッサの構成を具体的に論じようとする、何らかの対戦モデルを想定することが必要となる。本論文では、ある領域内に目標物が散在しており、2台のロボットが自律的に移動しながら競って目標を収集するという形式の対戦モデル^{(3),(4)}を設定する。本対戦ロボットを競争巡回ロボットという。

本論文では、次のような方法を提案する。

- ①先読み探索アルゴリズムとして、非交互着手 α - β 探索法⁽⁵⁾を採用する。
- ②先読み探索機能を可能な限りハードウェアによって実現する⁽⁶⁾。
- ③マルチプロセッサ構成とし、先読み探索を各プロセッサによって分担実行する。

非交互着手 α - β 探索法は、目標に直行する動作を単位とする方法であり、競争巡回ロボットの先読み探索を効率的に実行することができる。本方法は、非交互着手を取り扱うが、対戦

平成 14 年 12 月 26 日受理

* システム情報工学科・教授

** (株) シー・エス・デー

者は一般には非対称な機能を有すると考えられるので、本方法は汎用性の観点からも有効と考えられる。マルチプロセッサによる先読み探索の分担は、ある局面を与えてこれに対する最善着手を回答するという形式で、極めて簡潔、かつ有効に実行できると考えられる。

まず、非交互着手 α - β 法に基づいてプログラムの作成を行い、本方法の有効性を検証した。次に、非交互着手 α - β 法に基づいて、先読み探索プロセッサのハードウェア設計を行った。次いで、プロセッサの主要な回路を FPGA を用いて試作し、動作の検証を行った。

先読み探索速度に関する評価を行った結果、提案したプロセッサの速度は、汎用プロセッサを用いてソフトウェアによって処理する場合に比較して、単一プロセッサの場合で約 42 倍、マルチプロセッサとすると、さらにプロセッサの個数倍となることが分かった。

2. 対戦モデルの設定

本論文では、ある領域内に散在する目標物を競争しながら収集する対戦ロボット（競争巡回ロボットという）を例にとって検討する。対戦モデルを以下のように設定する。

〔ルール 1〕 ロボットの行動領域は $M \times N$ のマス目に分割される。

〔ルール 2〕 領域内に目標が散在する。障害物はない。

〔ルール 3〕 双方のロボットの機能は等しい。

〔ルール 4〕 ロボットの位置、および、目標の位置はマス目で表わされる。1 個のマス目には、一時に 1 個のロボットだけが存在できる。

〔ルール 5〕 ロボットは一つのタイミングで、現在位置の 8 近傍のマス目に移動できる。

〔ルール 6〕 2 つのロボットが同一のマス目に入ろうとするとき、いずれか一方が優先権を有する（優先権を有する側を先手、その反対側を後手という）。

〔ルール 7〕 目標の数は変数とする。

〔ルール 8〕 目標は価値点を有する。

〔ルール 9〕 目標の分布はランダムとする（1 回の対戦ごとに乱数によって決定する）。

〔ルール 10〕 ロボットが目標のあるマス目に入るとその目標の価値点が得点に加算される。そして、目標は削除される。

〔ルール 11〕 2 つロボットのスタート位置は、相互に対称な最も離れた 2 つのマス目にとる。

〔ルール 12〕 ある回数の対戦を行い、次の条件が成り立つときは先手の勝ち、そうでないときは後手の勝ちとする。

先手の得点 > 後手の得点 + H

ただし、H はハンデキャップを表わしており、多くの対戦結果に基づいて定める。

図 1 に、領域を 8×8 、目標の数 $B=6$ とした場合の初期局面の例を示している。図において、A は先手ロボット、B は後手ロボット、分数式 n/m は目標を表している。 n は目標の番号を、 m はその価値点を表わしている。

			B				
		3/1					
	2/1						6/1
				4/2			
						5/1	
	1/2						
				A			

図 1 初期局面の例

3. 先読み探索アルゴリズムとプログラム作成

3.1 先読み探索アルゴリズム

本対戦モデルでは、ロボットの動作の最小単位は1マスの移動となるが、この1マス単位の探索では、先読み探索に要する時間がかかり過ぎる。本研究者は、本対戦モデルに適した先読み探索方法として、目標に最短距離で直行する動作を単位とする方法(非交互着手 α - β 法という)を提案してきた⁽⁶⁾。本方法は、以下のように要約される。

まず、着手を次のように定義する。

[定義1] 着手: ある目標に向かって最短距離で直行することを着手という。

次の変数を設ける。

L: 探索の深さ

C: 2台のロボットの走行時間の差

M: 局面評価値の最大値

[アルゴリズム1] 非交互着手 α - β 法

(手順1) $L \leftarrow 0$, $C \leftarrow 0$, $M \leftarrow \infty$, 手番を先手にする。

(手順2) L =探索の深さの限界値、または、目標がなくなったときは、局面評価を行い、 $L \leftarrow L-1$, 手順5へ、さもなければ手順3へ。

(手順3) 着手候補をリストアップする。着手の対象とする目標は直面目標とする。ただし、相手がすでに同一目標に対して着手しており、かつ、相手の方が先に到着するときはリストに加えない。

(手順4) 一つの着手候補(優先順位の高いもの)を選択し着手する。

先手番ならば、 $C \leftarrow C + \text{目標までの所要時間}$

後手番ならば、 $C \leftarrow C - \text{目標までの所要時間}$

$C > 0$ ならば、手番を後手にする。

$C \leq 0$ ならば、手番を先手にする。

相手が同一目標に対してすでに着手しており、かつ、自分の方が先に到着できる場合は、相手側の着手を変更せしめる。

$L \leftarrow L+1$, 手順2へ。

(手順5) 先手番なら $M < \text{局面評価値}$, 後手

番なら $M < \text{局面評価値}$ のとき、 $M \leftarrow \text{局面評価値}$, かつ、最善着手候補 \leftarrow 着手。

着手を戻す。 C の値を戻す。手番を戻す。

(手順6) 未探索着手候補があれば手順4へ、さもなければ手順7へ。

(手順7) $L=0$ ならば終了、さもなければ、 $L \leftarrow L-1$, 手順4へ。

□

上記アルゴリズムにおいて、 L と C は全手順を通して単一の変数である。それ以外の変数は、すべて、再帰的に用いられる。

3.2 プログラム作成

アルゴリズム1に基づいてプログラム作成を行った。言語はC言語を用いた。完成したプログラムのステップ数は全部で、約400ステップとなった。

図2に、実行時の初期画面を示す。図において、井印はマス目の位置を、Aは先手ロボットの位置を、Bは後手ロボットの位置を、数字は目標の価値点を表わしている。

完成したプログラムについて、先読み探索に要する命令のステップ数を計算すると、目標の数 $B=15$, 先読みの深さを10手とすると、 4.6×10^8 ステップとなることが分かった。この値は、クロック周波数1GHzのコンピュータを用いた場合、約1.8秒に相当する。

```
#####A#####
#2#####
#####
#####
#####
#####
#####
#####
#1#####
#####B#9#
```

図2 実行初期画面

4. 先読み探索プロセッサの設計

4.1 プロセッサの全体構成

本論文では、いろいろな種類の対戦に共通に使用できる先読み探索プロセッサの実現を目標とする。具体的な対戦モデルが必要となる場合は、競争巡回ロボットを想定する。高速性を達成するため、可能な限りハードウェアによる並列処理を取り入れる。これは、次の2つに分かれる。

- ① 先読み機能をハードウェア化する。
- ② マルチプロセッサ構成とする。

本プロセッサを以下のように構成する。

〔構成1〕 プロセッサ全体は基本構成部（対戦のルールによって変わらない部分）と可変構成部（対戦のルールによって異なるハードウェアとなる）より成る。

〔構成2〕 基本構成部は、プログラム制御方式とする。機械語命令は先読み探索に必要な命令に特化する。

〔構成3〕 可変構成部は、可変論理構造とする。すなわち、FPGAの形式でプロセッサ内部に組み込む。

〔構成4〕 単位プロセッサを多数配列し、先読み探索を分担実行する。

図3に、プロセッサの全体構成を示す。

先読み探索の特徴より、マルチプロセッサによる仕事の分担は、局面と手番を与えこれに対する最善着手を回答するという形式で、極めて簡潔に行うことができる。このとき、一つのプロセッサは一つのエージェントととらえることができ、プロセッサ全体は、マルチエージェントシステムということができる。

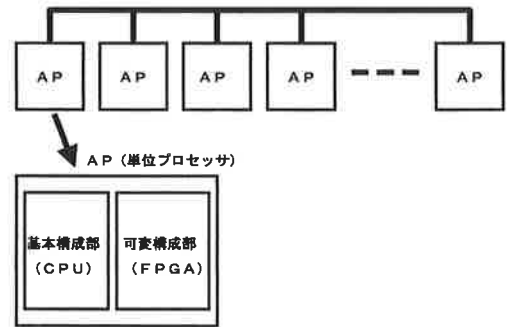


図3 プロセッサの全体構成

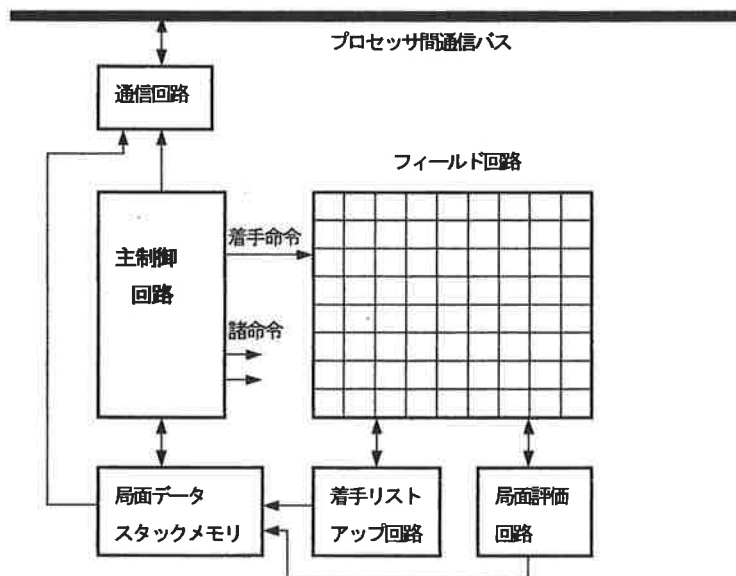


図4 単位プロセッサの全体構成

4.2 単位プロセッサの全体構成

プロセッサ1個（単位プロセッサという）は次の回路から成る。

- ① 主制御回路
- ② フィールド回路（ロボットの運動領域を模擬したハードウェア）
- ③ 着手回路
- ④ 着手リストアップ回路
- ⑤ 局面データ記憶スタックメモリ
- ⑥ 局面評価回路
- ⑦ 他のプロセッサとの通信回路

図4に、単位プロセッサの全体構成を示す。以下、各回路の詳細について論ずる。

4.3 主制御回路

本回路は、プログラム制御を実行する。プロセッサの命令構成を以下のように定める。

(1) 先読み探索命令

- ① 着手リストアップ命令
- ② 手番指定/交代命令
- ③ 1手着手命令
- ④ 1手戻し命令
- ⑤ 局面評価命令
- ⑥ 中途最良評価値、中途最良着手更新命令
- ⑦ ノード最良評価値、ノード最良着手登録命令

(2) 一般命令

- ① 加算命令
- ② 減算命令
- ③ データ移動命令
- ④ くり返し命令
- ⑤ 条件分岐命令
- ⑥ 入出力命令

主制御回路の構成は、先読み探索命令に関するデコード、その他の回路が付加される以外は、基本的には、汎用CPUの構成と同様となる。

4.4 フィールド回路

本回路は、ロボットの活動領域を模擬したハードウェアであり、先読み探索命令を高速に実行するよう構成する。

[構成5] マス目に対応するSRAM（マス目RAMという）を設ける。

[構成6] マス目RAMに、機械語命令を用いて所要データを書き込み/読み出しできるようにする。マス目の座標は機械語のオペランドにそのまま記述する形式とする。マス目RAMに書き込むデータは次のようになる。

- ① 目標が存在することを示す符号
- ② ロボットが存在することを示す符号
- ③ 後述する信号P、信号Qなど

[構成7] 各マス目の間に、次のような信号伝達スイッチとその制御回路を設ける。

(1) 信号Pの伝達

信号Pを受け取ったマス目は、その8近傍のマス目に信号Pを自動的に伝達する。従って、信号Pは次々に隣接するマス目に伝搬していく。

(2) 信号Qの伝達

信号Qを受け取ったマス目は、その3近傍、または、5近傍のマス目に信号Qを自動的に伝達する。3近傍には4通りの取り方(3近傍の方向という)が存在する。5近傍にも4通りの取り方が存在する。信号Qとともに、3近傍/5近傍の別と近傍の方向もデータとして一緒に伝達する。信号Qは次々に指定された方向のマス目に伝搬していく。

[構成8] 目標が存在するマス目は、信号Pが到着したとき次のような動作を行う。

① 信号Pを受信したことを示す信号（P受信信号という）をONにし、かつ、目標の番号をバスラインに出力する。P受信信号がすでに

ON になっているときは（他の目標が先に P 受信信号を出したことを示す）、OFF になるまで待つ。

② 信号 Q を隣接するマス目に発する。

〔構成 9〕 信号 Q を受信したマス目は、信号 P の受信とそのマス目以降への信号 P の伝達を停止する。

図 5 に、信号 P、信号 Q が伝搬していく様子を示す。目標が信号 P の発信元に対して 45 度の角度をなす位置にいるときは、信号 Q は 5 近傍に、それ以外の角度の場合は 3 近傍に伝搬する。

〔構成 10〕 各マス目は、目標が存在するか否かにかかわらず、次の動作を行う。ただし、信号 P はロボット A に対応する信号 P1 とロボット B に対応する信号 P2 の 2 種類を設ける。

① 信号 P1 が到着したマス目は信号 P2 の受信とそのマス目以降への伝達を停止する。

② 信号 P2 が到着したマス目は信号 P1 の受信とそのマス目以降への伝達を停止する。

構成 5、構成 6 は、着手回路の機能を果している。従って、着手回路として特別な回路を設ける必要はない。

構成 7、構成 8、構成 9 は着手リストアップに貢献する。

構成 10 は、局面評価に必要な 2 つのロボットの勢力圏を調べるのに貢献する。

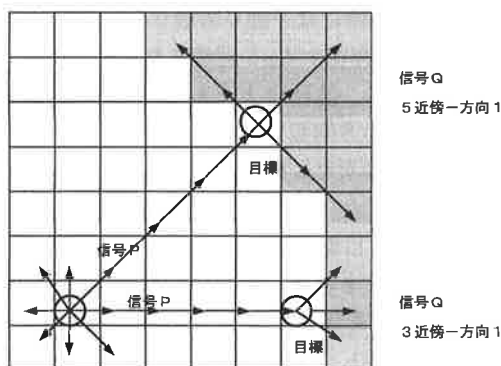


図5 フィールド回路における信号 P、信号 Q の伝搬

4.5 着手リストアップ回路

直面目標のリストアップは、フィールド回路の構成 7、構成 8、構成 9 を利用して次のように行う。

① 「着手リストアップ命令」は着手リストアップ回路を起動する。着手リストアップ回路は、まず、着手側ロボットの現在位置のマス目に信号 P を与える。

② 信号 P は自動的に 8 近傍に伝達していく。

③ 信号 P がある目標物に到着すると P 受信信号と目標の番号がバスラインに出力される。このとき、着手リストアップ回路は、自動的にその目標の番号を着手候補リストに登録する。登録が終了したら P 受信信号を OFF にする。信号 P の伝達が、複数の目標物に、同じタイミングで行われたときは、優先順位の高いマス目から順に P 受信信号を出力する。

④ 信号 P がある目標物に到着した後は、信号 Q がこの目標物以降の 3 近傍/5 近傍に伝達されて、この目標物の陰になる目標物がリストアップされないように働く。

⑤ 信号 P、または、信号 Q がフィールド全体に伝達され終わったとき、「着手リストアップ命令」の実行終了とする。

4.6 局面評価回路

局面評価は、フィールド回路の構成 10 を利用して次のように行う。

① 「局面評価命令」は局面評価回路を起動する。局面評価回路は、まず、着手側ロボットの現在位置のマス目に信号 P1 を、相手側ロボットの現在位置のマス目に信号 P2 を与える。

② 信号 P1、信号 P2 ともに、自動的に 8 近傍に伝達していく。

③ 信号 P1、信号 P2 は互いに相手がすでに伝達したマス目には伝達しないので、信号 P1 の伝達されたマス目がロボット A の勢力圏、信号 P2 が伝達されたマス目がロボット B の勢力

図と判定できる。

④ 局面評価回路は次の計算を行う。

評価点＝「ロボット A の既得点」－「ロボット B の既得点」＋「ロボット A の勢力圏にある目標物の価値点」－「ロボット B の勢力圏にある目標物の価値点」

⑤ 信号 P1, 信号 P2 がフィールド全体に伝達され終わったとき, 「局面評価命令」の実行終了とする。

4.7 局面データ記憶スタックメモリ

先読み探索を遂行する過程で, 次のような局面データ群を記憶する必要がある。

- ① 着手候補のリスト
- ② 着手候補の数
- ③ 探索中の着手の番号
- ④ 中途の最良着手候補
- ⑤ 中途の最良評価値
- ⑥ ノードの最善着手

これらのデータは, 探索の深さごとに (階層構造に) 記憶する必要がある。これらのデータの書き込み/読み出しについては, 最後に記憶された階層のものが最初に読み出される関係にあるので, その記憶にはスタックメモリが適している。

着手リストアップ回路で得られた着手リストや着手候補の数は, この局面データ記憶スタックメモリに直接 (プロセッサの命令によらずに) 送られるように回路を構成する。

4.8 マルチプロセッサシステム

単位プロセッサを並列接続し, 先読み探索を分担して実行する構成とする。

プロセッサ間の通信は次のように行う。

① 探索依頼を行うプロセッサはコール信号を ON にし, かつ, 自分の番号をバスラインに出力する。

② 手の空いているプロセッサはアクセプト信号を ON にする。

③ アクセプト信号が ON になったら, 依頼

元プロセッサは, バスラインを介して局面データを送る。

④ 探索を終了したプロセッサは, 探索終了信号を ON にし, かつ, 探索結果 (最善着手) を出力する。

なお, 各プロセッサは必ずしも単一の探索アルゴリズムに従う必要がなく, 複数の異なるアルゴリズムを実行することも可能である。

5. プロセッサの試作

単位プロセッサの以下の回路を FPGA を用いて試作した。① フィールド回路, ② 着手リストアップ回路, ③ 着手回路・局面データスタックメモリ, ④ 局面評価回路, ⑤ 通信回路。

FPGA は, ザイリンクス社製 XC9536 キットを用いた。図 6 に, そのキットを示す。

各回路が, いずれも, 正常に機能することを確認した。

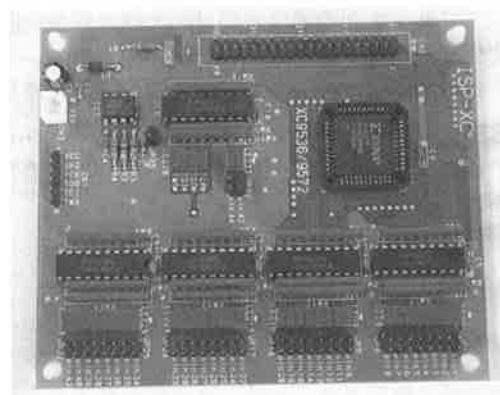


図6 FPGA キット

6. 評価

本論文で提案した先読み探索プロセッサの評価を試みる。比較対象として, 「汎用マイクロプロセッサを用いてソフトウェアによって処理する方法」をとる。比較パラメータとして処理速度をとるが, 処理速度の絶対値は明らかにしにくいので, 先読み探索に要する命令のステップ

数を比較することとする。競争巡回ロボットについて、目標の数、先読み探索の深さなどの条件を指定して、最善着手を得るのに要する命令のステップ数を数え上げる。

(1) 単一プロセッサの比較

提案したプロセッサのステップ数を S_1 、同じ条件で汎用プロセッサを用いた場合のステップ数を S_2 と表わす。

目標数を 15 個、先読みの深さを 10 手とすると、次の値が得られる。

$$S_1 = 1.1 \times 10^7 \text{ ステップ}$$

$$S_2 = 4.6 \times 10^8 \text{ ステップ}$$

よって、次が得られる。

[結論 1] 提案したプロセッサの先読み探索速度は、単一プロセッサの場合、汎用プロセッサのソフトウェアによる場合に比較して、約 42 倍となる。

(2) マルチプロセッサの比較

プロセッサ間の通信時間は、先読み探索の時間に比較して十分小さいので次が得られる。

[結論 2] 提案したプロセッサの先読み探索速度は、マルチプロセッサの場合、プロセッサの数を M とすると、単一プロセッサの M 倍となる。

7. む す び

本論文では、いろいろな種類の対戦ロボットに共通に使用でき、かつ、高速性を有する先読み探索プロセッサの構成法を検討した。対戦モデルとして、ある領域内に散在する目標物を競って収集するという形式の対戦を想定し、次のような構成法を提案した。

① 先読み探索アルゴリズムとして、非交互着手 α - β 探索法を採用する。

② 先読み探索機能を可能な限りハードウェアによって実現する。

③ マルチプロセッサ構成とし、先読み探索を各プロセッサで分担実行する。

以下の手順でプロセッサの設計・製作を行い、正常に機能することを確認した。

① 非交互着手 α - β 法に基づいてプログラムを作成した。

② 先読み探索プロセッサのハードウェア設計を行った。

③ プロセッサの主要な回路を FPGA を用いて試作した。

先読み探索速度の評価を行った結果、提案したプロセッサによれば、汎用プロセッサを用いてソフトウェアによって処理する場合に比較して、単一プロセッサの場合、約 42 倍、マルチプロセッサとすると、さらにプロセッサの個数倍となることが分かった。

今後、次のような課題について検討する予定である。

① 本プロセッサに適した FPGA 構成。

② マルチプロセッサへの探索割付方法。

③ LSI チップの試作。

文 献

- (1) 人工知能学会編、人工知能ハンドブック、オーム社、1990.
- (2) 小谷、吉川他、コンピュータ将棋、サイエンス社、1990.
- (3) 苫米地宣裕、“競争巡回形知能ロボット制御用先読みエンジンの設計”，計測自動制御学会東北支部研究集会資料，175-4，1998-6.
- (4) 苫米地宣裕、藤枝孝徳、“競争巡回ロボット制御用先読み探索プロセッサの構成法”，計測自動制御学会東北支部 35 周年記念学術講演会予稿集，pp. 99-100，1999-10.
- (5) 苫米地宣裕，“競争巡回知能ロボットとその先読み探索アルゴリズム”，八戸工業大学情報システム工学研究所紀要，vol. 12，pp. 1-6，2000-3.
- (6) 藤枝孝徳，八戸工業大学大学院修士学位論文：対戦ロボット制御用先読み探索プロセッサの構成法，2001-3.